DRAFT
DO NOT CITE OR QUOTE

NCEA-F-0644
July 1999
Review Draft

Guidelines for Carcinogen Risk Assessment

NOTICE

THIS DOCUMENT IS A PRELIMINARY DRAFT. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent Agency policy. It is being circulated for comment on its technical accuracy and policy implications.

Risk Assessment Forum
U.S. Environmental Protection Agency
Washington, DC

DISCLAIMER

The current document constitutes work in progress. It incorporates some changes to the January 1999 review draft based on discussions at the January meeting and the recently released draft letter from the Science Advisory Board (SAB), dated May 20, 1999. The Agency is continuing to address the SAB recommendations. However, for the purpose of providing a context for a discussion of the guidance on assessing children's risk, the Agency has provided the most current version of the draft guidelines.

The document is a draft for review purposes only. It does not constitute U.S. Environmental Protection Agency policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

GUIDELINES FOR
CARCINOGEN RISK ASSESSMENT
FRL-

AGENCY: U.S. Environmental Protection Agency

ACTION: Notice of Final Guidelines for Carcinogen Risk Assessment

SUMMARY: The U.S. Environmental Protection Agency (EPA) is today publishing a document entitled *Guidelines for Carcinogen Risk Assessment* (hereafter "Guidelines"). These guidelines were developed as part of an interoffice guidelines development program by a Technical Panel of the Risk Assessment Forum within EPA's Office of Research and Development. These guidelines revise and replace EPA's 1986 *Guidelines for Carcinogen Risk Assessment* published on September 24, 1986 (51 FR 33992).

In an associated *Federal Register* notice, the Agency discusses its cancer assessment prioritization process, including priorities for reassessments under these final guidelines.

FOR FURTHER INFORMATION, CONTACT: Technical Information Staff, Operations and Support Group, National Center for Environmental Assessment-Washington Office, telephone: 202-564-3261.

ADDRESSES:

The Guidelines will be made available in the following ways:

1) The electronic version will be accessible on EPA's Office of Research and Development home page on the Internet at http://www.epa.gov/ORD

2)

3) This notice contains the full document. In addition, copies will be available for inspection at EPA headquarters and regional libraries, through the U.S. Government Depository Library program, and for purchase from the National Technical Information Service (NTIS), Springfield, VA; telephone: 703-487-4650, fax: 703-321-8547. Please provide the NTIS PB No. () ($xx.00) when ordering.

SUPPLEMENTARY INFORMATION: In 1983, the National Academy of Sciences (NAS)/National Research Council (NRC) published its report entitled *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983). In that report, the NRC recommended that Federal regulatory agencies establish "inference guidelines" to ensure consistency and technical quality in risk assessments and to ensure that the risk assessment process was maintained as a scientific effort separate from risk management. The 1986 cancer guidelines were issued on September 24, 1986 (51 FR 33992). The Guidelines published today continue the guidelines development process. These guidelines set forth principles and procedures to guide EPA scientists in the conduct of Agency cancer risk assessments and to inform Agency decision makers and the public about these procedures.

These guidelines contain inference guidance in the form of default inferences to bridge gaps in knowledge and data. Research conducted in the past decade has elucidated much about the nature of carcinogenic processes and continues to provide new information. These guidelines take account of knowledge available now and to provide flexibility for the future in assessing data and employing default inferences, recognizing that the guidelines cannot always anticipate future research findings. Because methods and knowledge are expected to change more rapidly than guidelines can practicably be revised, the Agency will update specific assessment procedures with peer-reviewed supplementary, technical documents as needed. Further revision of the guidelines themselves will take place when extensive changes are necessary.

Since 1986, the EPA has sponsored several workshops about revising the cancer guidelines (U.S. EPA, 1989b, 1989c, 1994a). The Society for Risk Analysis conducted a workshop on the subject in connection with its 1992 annual meeting (Anderson et al., 1993). Participants in the most recent workshop in 1994 reviewed an earlier version of the guidelines proposed here and made numerous recommendations about individual issues as well as broad recommendations about explanations and perspectives that should be added. Many persons commented on the proposal of these guidelines in 1996, and all of these comments were considered. The EPA appreciates the efforts of all commenters and participants in the process and has considered their recommendations and concerns. An overview of the major features of the guidelines is provided below, followed by responses to comments on major science and science policy issues.

Overview of Major Features of the Guidelines

Characterizations

The guidelines call for greater emphasis on characterization discussions for hazard, dose response, and exposure assessment. These discussions will summarize the assessments to explain the extent and weight of evidence, major points of interpretation and rationale for their selection, and strengths and weaknesses of the evidence and the analysis, and to discuss alternative conclusions and uncertainties that deserve serious consideration (U.S. EPA, 1995). They serve as starting materials for the overall risk characterization process which completes the risk assessment.

Weighing Evidence of Hazard

The guidelines emphasize the weighing of all of the evidence in reaching conclusions about the human carcinogenic potential of agents. This is to be accomplished in a single step after assessing all of the individual lines of evidence. This is in contrast to the step-wise approach which was called for in the 1986 guideline. Evidence to be considered include tumor findings in humans and laboratory animals, an agent's chemical and physical properties, its structure-activity relationships to other carcinogenic agents, and its activities in studies of carcinogenic processes. Data from human studies are preferred for characterizing human cancer hazard. However, all of the above-mentioned information could provide valuable insights into its possible mode(s) of action and likelihood of human cancer hazard and risk. The guidelines recognize the growing sophistication of research methods, particularly in their ability to reveal the modes of action of carcinogenic agents at cellular and subcellular levels as well as toxicokinetic processes. The term mode of action is defined as a series of key events and processes starting with interaction of an agent with a cell, and proceeding through operational and anatomical changes resulting in cancer formation. "Mode" of action is contrasted with "mechanism" of action, which implies a more detailed understanding and description of events, often at the molecular level, than is meant by mode of action.

Weighing of the evidence includes addressing not only the likelihood of human carcinogenic effects of the agent but also the conditions under which such effects may be expressed, to the extent that these are revealed in the toxicological and other biologically important features of the agent.

Weight of Evidence Narrative and Hazard Descriptors

The weight of evidence narrative to characterize hazard summarizes the results of the hazard assessment and provides a conclusion with regard to human carcinogenic potential. The narrative explains the kinds of evidence available and how they fit together in drawing conclusions, and points out significant issues/strengths/limitations of the data and conclusions. As the narrative also summarizes the mode of action information, it also sets the stage for the discussion of the rationale underlying a recommended approach to dose response assessment.

In order to provide some measure of consistency in an otherwise free-form, narrative characterization, standard descriptors are utilized as part of the hazard narrative to express the conclusion regarding the weight of evidence for carcinogenic hazard potential. There are five standard hazard descriptors: *"carcinogenic to humans"*, *"likely to be carcinogenic to humans"*, *"suggestive evidence of carcinogenicity but not sufficient to assess human carcinogenic potential"*, *"data are inadequate for an assessment of human carcinogenic potential"*, and *"not likely to be carcinogenic to humans"*. Each standard descriptor may be applicable to a wide variety of data sets and weights of evidence and are presented only in the context of a weight of evidence narrative. Furthermore, more than one conclusion may be reached for an agent. For instance, using a descriptor in context, a narrative could say that an agent is *likely* to be carcinogenic by inhalation exposure and *not likely* to be carcinogenic by oral exposure.

Mode of Action

The use of mode of carcinogenic action in the assessment of potential carcinogens is the main thrust of these guidelines. This area of emphasis arose because of the significant scientific breakthroughs that have developed concerning the causes of cancer induction. In the absence of mode of action information, EPA takes conservative (public health protective) default positions regarding the interpretation of toxicologic and epidemiologic data. Animal tumor findings are judged to be relevant to humans, and cancer risks are assumed to conform with low dose linearity. Elucidation of a mode of action for a particular cancer response in animals or humans is a data rich determination. Significant information must be developed to ensure that a mode of action underlies the process leading to cancer at a given site. Understanding of mode of action can be a key to identifying processes that may cause chemical exposures to differentially affect children or males or females or other segments of the overall population. Some modes of action are

anticipated to be mutagenic and assessed with a linear approach for most, if not all, parts of the population. This is the mode of action of radiation and several other agents which have been recognized as known carcinogens. Several mutagenic carcinogens are also in utero carcinogens. Other modes of action may be assessed with either linear or nonlinear approaches only after a rigorous analysis of available data in accordance to the guidance as provided in the framework of mode of action analysis.

Dose-Response Assessment

Dose-response assessment evaluates potential risks to humans at particular exposure levels. The approach to dose-response assessment for a particular agent is based on the conclusion reached as to its potential mode(s) of action for each tumor type. Since an agent may induce multiple tumor types, the dose-response assessment includes an analysis of all tumor types, followed by an overall synthesis which includes the consistency of risk estimates across tumor types, the strength of the mode of action information of each tumor type, the anticipated relevance of each tumor type to humans including sensitive subpopulations (e.g., children).

Dose-response assessment for each tumor type is performed in two steps- assessment of observed data to derive a point of departure, followed by extrapolation to lower exposures, to the extent that is necessary. Dose-response data from human studies are preferred for estimating risks. When animal studies are the basis of the analysis, the estimation of a human equivalent dose utilizes toxicokinetic data to inform cross-species dose scaling, if appropriate and adequate data are available. Otherwise, default procedures are applied. For oral dose, the default is to scale daily applied doses experienced for a lifetime in proportion to body weight raised to the 0.75 power. For inhalation dose, the default methodology estimates respiratory deposition of particles and gases and estimates internal doses of gases with different absorption characteristics. Guidance is also provided for adjustment of dose from adults to children.

Response data on effects of the agent on carcinogenic processes are analyzed (non-tumor data) in addition to data on tumor incidence. If appropriate, the analyses of data on tumor incidence and on precursor effects may be combined, using precursor data to extend the dose response curve below the tumor data. Even if combining data is not appropriate, study of the dose response for effects believed to be part of the carcinogenic process influenced by the agent may assist in evaluating the relationship of exposure and response in the range of observation and

at exposure levels below the range of observation.

The first step of dose- response assessment is evaluation within the range of observation. Approaches to analysis of the range of observation of human studies is determined by the type of study and how dose and response are measured in the study. In the absence of adequate human data for dose response analysis, animal data will generally be used. If there are sufficient quantitative data and adequate understanding of the carcinogenic process, a biologically-based model may be developed to relate dose and response data on an agent-specific basis. Otherwise, as a default procedure, a standard model is used to curve-fit the data. The lower 95% confidence limit on a dose associated with an estimated 10% increased tumor or relevant nontumor response (LED_{10}) is identified. This generally serves as the point of departure for extrapolating the relationship to environmental exposure levels of interest when the latter are outside the range of observed data. Other points of departure may be more appropriate for certain data sets; as described in the guidance, these may be used instead of the LED_{10}. The LED_{10}, rather than the ED_{10} (the estimate of a 10% increased response), is the proposed standard point of departure for two reasons. One is to permit easier comparison with the benchmark dose procedure for noncancer health assessment—also based on the lower limit on dose. Another is that the lower limit, as opposed to the central estimate, accounts for the variability in the experimental data.

The second step of dose-response assessment is extrapolation to lower dose levels, if needed. This is based on extension of a biologically based model if supportable by substantial data. Otherwise, default approaches are applied that are consistent with current understanding of mode(s) of action of the agent. These include approaches that assume linearity or nonlinearity of the dose response relationship, or both. The default approach for linearity is to extend a straight line to zero dose/ zero response. The linear approach is used when there is an absence of sufficient information on modes of action, or the mode of action information indicates that the dose-response curve at low dose is or expected to be linear. The default approach for nonlinearity is to use a margin of exposure analysis rather than estimating the probability of effects at low doses. The use of a margin of exposure approach is included as a new default procedure to accommodate cases in which there is sufficient evidence of a nonlinear dose response, but not enough evidence to construct a mathematical model for the relationship. A margin of exposure compares the point of departure (e.g., LED_{10} of a non-tumor response, a key event necessary for the carcinogenic process) with the dose associated with the environmental exposure(s) of interest by computing the ratio between the two. The margin of exposure analysis explains the biological

considerations for comparing the observed data with the environmental exposure levels of interest and assists the decision-maker to determine an acceptable level of exposure in accordance with requirements of the statute under which the risk management decision is being made. There are several factors to be considered. These include: the nature of the response (e.g., tumors or precursor events) used in the dose response assessment, the slope of the dose response curve at the point of departure (e.g., shallow or steep), human sensitivity to the response as compared with laboratory animals (if animal data are used), the nature and extent of human variability to the response (e.g., children versus adults, men versus women), and the nature of anticipated human exposure and characteristics of the populations potentially at risk. If, in a particular case, the evidence indicates a biological threshold, as in the case of carcinogenicity being secondary to a toxicity that has a threshold, a reference dose (RfD) or reference concentration (RfC) like approach may be considered. The RfD/RfC approach would include determination of a point of departure (e.g. LED_{10}) for a precursor event on the path toward carcinogenesis, and application of uncertainty factors for cross-species and inter-individual variation and perhaps, others as needed. In this case, the RfD or RfC would be an estimate with uncertainty spanning perhaps an order of magnitude of daily exposure to the human population (including sensitive subgroups) that is anticipated to be without cancer hazard despite a lifetime of exposure.

Response to Comments on Major Issues

The EPA Science Advisory Board reviewed these guidelines at the proposal stage, and has provided further advisory review of specific issues. In its proposal, the EPA asked for comment on several major issues. These are discussed below. Other comments on science issues are discussed in Appendix G.

Use of Mode of Action Information

The proposal to use mode of action information in both characterizing hazard potential and in estimating dose-response was generally endorsed by commenters as appropriate for employing new scientific knowledge and in opening the potential for consideration of alternative tests of carcinogenesis (e.g., in genetically engineered animals), as well as new approaches to collecting human data (e.g., molecular epidemiology.) The SAB (SAB, 1997) noted:

> "The EHC (Environmental Health Committee) generally endorsed the Guidelines' mode of action proposals, but suggested that the Guidelines contain specific criteria for judging that the data on mode of action are valid and adequate."

There was a general call by public commenters for more guidance and examples of how to decide that mode of action data on an agent are adequate to use.

These final guidelines continue the direction of the proposal as to use of mode of action data. In response to SAB and public views, they include a framework for evaluation of data regarding a postulated mode of action, and case examples to illustrate how the framework is used. The framework can be used for assessments of potential modes of action whether the endpoint of interest is cancer or other toxicity. It is also used to consider whether there is special concern for certain subpopulations (e.g., children), or for those who may have important metabolic differences. It should also be noted that the 1996 Food Quality Protection Act calls for assessment and regulatory consideration of the combined exposure to food-use pesticides that have a common mechanism of action. This is an application of mode of action data that follows existing EPA practice with regard to certain classes of environmental chemicals and pesticides.

Children's Health

Some public commenters urged the EPA Administrator not to accept or finalize the guidelines until specific provisions and/or guidance are incorporated that will ensure protection of the fetus, infants and children. Some public commenters also opined that the nonlinear dose-response approach is less protective for children.

On the question of protecting children, the final guidelines contain more guidance to support examination for potentially disproportionate hazard and risk to children and other sensitive subpopulations related to the hazard and mode of action evaluations, dose-response and exposure assessments, and risk characterization. In particular, the framework for mode of action analysis requires an evaluation of whether a postulated mode of action is relevant to humans and is qualitatively applicable to subpopulations of concern including children. Ideally we would have data pertinent to the question with respect to the agent under assessment. In the absence of such data, a cogent biological rationale needs to be developed regarding whether the mode of action is applicable to children. This evaluation would include data on chemicals that are structurally related to the agent under review, and other scientific information at large, including such considerations as age-related similarities and differences in the occurrence of the specific tumor in the population, in the uptake, metabolism, and excretion of the agent, and in the occurrence of pertinent biochemical, physiological, and toxicological processes, including key events associated with the mode of action. Illustrations of such evaluation are provided in two case examples in Appendix D (chemicals T and Z).

A review of available information on adult and childhood exposures to radiation, pharmaceuticals and viruses indicate a similar spectrum of tumors in children and adults. Similar tumors were also observed in animal studies of a limited number of chemicals following combined perinatal and adult exposure compared to adult exposure alone. Therefore, it could be inferred that, from a qualitative standpoint, the same kind of tumors may develop following childhood or adult exposure to environmental chemicals. However, to ensure protection of children, these guidelines take the following default positions: when there are no agent-specific data or there is not a cogent rationale supporting the comparability between responses in children and adults, the postulated mode of action for producing tumors in adults is not considered operative in children, and a linear dose-response relationship will be used for the general population, including sensitive subpopulations, or more specifically, to infer potential risks for children as a default procedure. It

should also be noted that from a quantitative perspective, age differences in toxicokinetics and exposure may lead to greater or lesser occurrence of key events. Such differences may need separate evaluation and result in separate risk estimates for the young or for that portion of a lifetime. A case example for dose-response assessment is given to show how data can be used when early life exposure is associated with increased tumorigenesis (Appendix F).

The following review of current knowledge to address the question, "are cancers in children and adults the same" forms the basis for these default positions.

Are cancers in children and adults the same?

With notable exceptions, cancer is a set of diseases of advancing age. Childhood, the first 15 years of life, includes about 28% of the population and constitutes about 20% of a lifetime. Only about 0.3% of human cancers occur in childhood. The incidence of childhood cancer is about 15 per 100,000 individuals; the rate does not vary much among countries, usually differing by less than 2-fold worldwide (Parkin et al., 1988). Cancer incidence rates increased about 10-40% in the U.S. during the period between the 1970s-1990s, while the mortality rates decreased about 40%, largely a reflection of improved treatment for some childhood cancers (Gloeckler Ries et al., 1996).

Causes of human cancers appear to vary with the tumor type. Factors so far identified include inherited conditions (Tomlinson, 1997), associations with congenital malformations (Bosland, 1996; Cortes, 1998) and a variety of biological, physical and chemical factors. In some cases tumors in children and adults have been compared. Children and adults not uncommonly develop the same spectrum of tumors when they have inherited gene and chromosomal mutations, like Li-Fraumeni syndrome. With ionizing radiation which operates through mutagenic means, both the young and the old develop the same tumors, with the only difference generally being that children are about 2-fold more sensitive (NRC, 1990). Studies with anticancer drugs (cytotoxic and immunosuppressive) demonstrate again a similar spectrum of tumors (Hale et al., 1999; Kushner et al., 1998; Larson et al., 1996; Nyandoto et al., 1998). Various viral infections like Epstein Barr and hepatitis B lead to lymphoma and liver cancer, respectively, in both age groups (Lindahl et al., 1974; Mahoney, 1999). These observations indicate that the mode of action for these agents would be the same for children and adults.

Additional support to this finding comes from results of a review of about 40 rodent carcinogenicity studies with combined perinatal and adult chronic chemical exposure and adult chronic exposure alone. Although the data base is not robust, the findings were consistent across the studies and lead to three conclusions: perinatal exposure seldom produced types of tumors

not found in the standard bioassay, exposure beginning during the neonatal period and continuing throughtout the animal's lifetime often produced a higher incidence of tumors, and the latency period to tumor occurrence appeared to decrease with a perinatal exposure component (U.S. EPA, 1996). There are several consequences of these findings. It would seem that the developing organism is sensitive to the carcinogenic potentialities of some but certainly not all chemical agents. When cancer is induced following perinatal exposure, the sites are like those seen after extra utero exposure alone, with few exceptions. These exceptions will be addressed separately. Likewise, cancer incidence may be greater following combined exposure than with extra utero exposure alone. The nature of this difference needs further study. On the one hand, the young may be innately more sensitive to carcinogenic effects; on the other hand, the increased incidence may reflect a greater total dose or a greater time of dosing or possibly a combination of factors.

Most often differences between carcinogenic effects in the young vs. adults can be traced to differences in the handling of chemical agents. The fetus, infant and child may have metabolic capabilities that are qualitatively or quantitatively at variance with those in adults. The young may lack the capability to handle an exogenous chemical, which can have variable effects. If the parent compound is the toxic moiety, children may have enhanced susceptibility compared to adults, whereas they may have less hazard potential when it is a metabolite which has toxic properties (Snodgrass, 1992). Quantitatively, metabolism in the young is often faster than in adults which can also lead to corresponding changes in dose and, thus, cancer risk (Renwick, 1998).

In experimental animal studies focusing on perinatal exposure only, positive responses are noted only for strong mutagenic compounds that are positive in multiple species in traditional chronic bioassays on adults (Flammang et al., 1997). It would seem that if significant advances in detecting in utero and early extra utero environmental carcinogenic influences in rodents are to be made, they must await development of new understanding and technologies. Certain rodent transgenic systems may be rewarding.

Information on unique carcinogenic effects following in utero (or early postnatal) exposure is very unusual indeed. In humans, pharmacological use of diethylstilbestrol (DES) during pregnancy for threatened abortion resulted in increased incidence of clear cell adenocarcinoma of the vagina but not of other cancers in the daughters exposed in utero. It is thought that DES induces malformations during development that put vaginal cells at risk for cancer. For instance, risks for vaginal adenosis and dysplasia are high, while cancer only develops in about 1.5 in 1000 exposed persons (Hatch et al., 1998; Robboy et al., 1984; Vessey, 1989). Mice treated neonatally with tamoxifen develop uterine carcinoma, while animals dosed as adults

are free of cancer; humans develop uterine carcinomas from tamoxifen (Newbold et al., 1997; Wogan, 1997). In rats, chronic exposure with saccharin that commences in utero or early extra utero results in development of bladder cancer, while initiation of exposure at postweaning does not (Cohen and Ellwein, 1991); the reason for this difference is not fully understood.

Although there are similarities between childhood and adult tumors, significant differences exist. Tumors of childhood generally show more embryonic cell tumors, while adults have more carcinomas. Some tumors are quite unique to the young like, tumors of the sympathetic nervous system or adrenal medulla (neuroblastoma), kidney (Wilm's tumor), eye (retinoblastoma) or bone (Ewings' sarcoma). Such findings suggest that the carcinogenic process inhibits normal cell differentiation or enhances dedifferentiation to an embryonal cell type. Unlike many adult cancers, rarely has it been possible to identify environmental causes of childhood cancer, attributable to some degree to the rarity of most childhood cancers. Many of the childhood cancers occur as manifestation of hereditary syndromes; inherited factors are less prominent among adult cancers. Finally, there is often a very restricted number of gene and chromosomal mutations of cellular control factors in childhood cancers, whereas there are many different changes in adult cancers (Grufferman, 1998; Israel, 1995). Animal models for most of the tumors in childhood do not exist. More work is needed to discern the modes of action of these rare tumors and to understand the potential role of environmental influences.

Inter-individual Variation

The 1996 Proposed Guidelines called for a factor of no less than 10 as a default value for human variability in the analysis of margin of exposure, when application of the framework of mode of action analysis supports a nonlinear dose response relationship at low doses. The guidelines have not adopted application of an additional uncertainty factor when dose-response is considered by a linear extrapolation procedure.

Some public commenters questioned the validity and adequacy of the proposed default value of 10 (in the absence of agent specific data) to account for inter-individual variation in the margin of exposure analysis. The concern was that it would not be protective of fetuses, infants, or children.

The SAB and some public commenters also called for more attention in the guidelines and in further investigation to the subject of inter-individual variation. Some commenters particularly asked for reconsideration of the need for adding an uncertainty factor to the results of linear dose-response assessment, citing the NRC, 1994 report, "Science and Judgment in Risk Assessment"

which suggested adoption of such a factor with respect to estimates of individual risk. In particular, some have been concerned that use of mode of action considerations to depart from the linear dose-response approach might be insufficiently protective of fetuses, infants or children. The NRC recommendation was discussed in the 1996 proposal of these guidelines, but is more extensively discussed here.

The NRC, 1994 report recommended that EPA adopt a default assumption as to inter-individual variability in susceptibility among humans to be used as an added factor of conservatism with linear dose-response extrapolation. The report discussion indicated that this was not relevant for estimating risk to the general population unless bias exists in an estimate of average risk, but would apply to estimating individual risk (p. 208). The report was commissioned under the 1990 Clean Air Act, and the context of discussion was the estimation of risk from air emission sources. "Population risk" in this context means all who are subject to a particular kind of emission in the U.S. "Individual risk" means those who are in special circumstances such as being the most exposed to an emission plume. The latter term may also apply to other special exposure scenarios such as local exposures associated with waste sites or use of a product. The report regarded the EPA linear risk extrapolation estimation as one of "average" or "median" risk because generally rodent study results are used (with inter-species dose adjustment) to estimate human risk without including adjustment for human variability. This results, in the report's view, in extrapolation of response data from average rodent to average human. The report reviewed studies estimating potential variation in susceptibility to cancer and found them to generally indicate that "predisposed" people may be a factor of 10 more susceptible than "normal" ones. The studies were in part of cancer mortality data that would cover all sources of variability including, diet, personal habits, nutrition, inborn metabolism factors and genetic disease, infections, exposure to radiation or chemicals, medical care, and all sources of exposure.

The Committee was divided on the question of whether an explicit factor of 10 in EPA assessments of individual risk was justified at this time. The studies reviewed were summarized as giving a first approximation of variability from all sources as a lognormal distribution with about 5% of the population at the two ends of the distribution being 25 times more or less susceptible than average, and 1% at the extremes being 100 times more or less susceptible than average. The report recommended an extensive list of research by the Federal government to inquire into the extent of variability and the factors involved, and for examination of the adequacy of the 10-fold factor traditionally used in non cancer toxicity assessment. Examples of factors conferring susceptibility to cancer were discussed in Appendix H-2 to the report. These included genetic diseases that lead to inability to repair DNA damage or are mutations in tumor suppressor genes

(Li-Fraumeni, retinoblastoma, familial polyposis coli), viral diseases (hepatitis B, Epstein-Barr), immune deficiency, nutritional factors, as well as toxicokinetic differences.

Reasoning about the following two questions provides the bases for the positions taken in these guidelines and the responses to the SAB and public comments on these two issues--In the absence of agent-specific data, is a default factor of 10 for inter-individual differences adequate, given current data? Is the overall result of using linear extrapolation adequately protective without an additional factor?

In the absence of agent-specific data, is a factor of 10 for inter-individual differences adequate for carcinogenic response, given current knowledge?

Available information indicates the use of a 10-fold factor, as a default, is appropriate. The incidence of cancer at given sites among human beings does not vary widely among persons with very different diets in the U.S. and worldwide. Generally, the variation is within 10-fold (Schottenfeld and Fraumeni, 1996; Parkin et al., 1988; NRC, 1990; ICRP, 1991) It is apparent that cancer should be considered as a product of the balance of many risk and protective factors, as is the familiar approach with heart and other diseases. Diet, genetic background, infectious disease, lifestyle and other circumstances for individuals all are known to influence cancer risk.

It is also appropriate to consider data that examine human variability for noncancer effects since nontumor response (i.e., a precursor response) is generally used in the margin of exposure approach. EPA has traditionally used an uncertainty factor of 10 for human variability in susceptibility in assessment of noncancer endpoints. The factor of 10 has been considered to cover two elements of uncertainty in the absence of data: toxicokinetics (processes which determine delivery of the active agent to the site of activity) and toxicodynamics (processes which determine the extent of response) (U.S. EPA, 1994c; Renwick, 1993, 1997, 1998; WHO, 1994).

Several studies have examined variability among humans and some have examined the coverage and adequacy of the traditional 10-fold factor (Renwick, 1993, 1997, 1998; Calabrese, 1985; Dourson and Stara, 1983; NRC, 1993b). Most of the human effect data have been about therapeutic drugs or non cancer effects of various chemicals. The results inform the issue of human variability generally. The data represented a range of metabolic or clearance pathways for the agents included that would apply as well to xenobiotics. These studies support a conclusion that the 10-fold factor is adequate as a default, but important exceptions need to be addressed in individual cases, such as a case in which there is polymorphism in a major pathway of elimination (e.g., enzyme kinetics), particularly if the pathway results in detoxification and if there is no

compensation through alternative pathways. This is an area for further research and for analysis of existing data and generation of guidance on pathways for major structural groups of xenobiotics.

Renwick (1998) also analyzed therapeutic drug literature to study whether toxicokinetics in infants and children indicate the need for an increased uncertainty factor and compared human with rodent data. The examination included consideration and consequences associated with the time it takes for major toxicokinetic functions to achieve adult competency including absorption, distribution, and elimination mechanisms for xenobiotics as well as data on these functions at different ages for a number of drugs. The analysis concluded that an increased uncertainty factor for toxicokinetics for post-suckling infants and children is not required. Moreover, the higher clearance of many xenobiotics by children may compensate, at least in part, for potential increased organ sensitivity during development. In addition, the calculation of dose on a body weight basis can provide an extra margin of safety. This approach to dose calculation has been adopted in these guidelines. Exceptions that require agent-specific analysis of the adequacy of the default are an exposure scenario that applies to neonates and involves a pathway that is not mature, such as cytochromes P450 whose maturation has been studied (Cresteil, 1998) or an agent that may be more rapidly activated by children.

While there are exceptions that require attention for particular agents, available studies support the use of the 10-fold factor, as a default. It should be noted that pre-existing disease or genetic constitution may place a percentage of the population at special risk and the factor of 10 would not cover this. The Agency intends to develop supplementary guidance for incorporating these considerations in the margin of exposure analysis.

Is the overall result of using linear extrapolation adequately protective, without an additional factor?

The question is whether a linear extrapolation of tumor data from animal studies to estimate individual risk should incorporate an extra factor for human variability. Given that toxicokinetic or scaling adjustment is made to the animal dose to derive an human equivalent dose (HED), the remaining question is whether the inherent conservatism of linear extrapolation is adequate to cover uncertainties regarding human variation. Considerations not included in the NRC 1994 report are pertinent.

First, dose-response assessments conducted for most environmental agents rely on data from chronic animals bioassays which are conducted at high doses, with the highest dose selected to produce some toxicity, a maximum tolerated dose. This is done to increase detection of any carcinogenic property of an agent qualitatively. The consequence is that toxicokinetic, stress, and other parameters may not be representative of those at low dose. The "average" rodent has thus been tested under sensitizing conditions and may not be appropriately characterized as "average" in the sense of the NRC report. The extrapolation from such bioassays is more appropriately described as from "sensitive", not "average" animals. Second, EPA (U.S. EPA, 1986b) has always described the linear procedure using the linearized multistage (LMS) procedure as resulting in a plausible upper bound estimate of risk at low dose where true risk may be lower, including zero. The linear procedure adopted under these guidelines has the same characterization and yields low dose risk estimates that are close to those of the LMS procedure. On theoretical grounds (Lutz, 1990b) as well as from study of animal data sets (ILSI, 1995), a linear extrapolation is generally a conservative approach that is believed, with few exceptions, to over-estimate risk at low doses to varying degrees. This is important in the context of the traditional low-dose target range of EPA risk management programs: 10^{-6} to 10^{-4} risk for population and individual risk. This is a range that is usually several orders of magnitude below the range of observed data in animal bioassays and the range of observation in most human studies and has been the range considered protective of the general and sensitive populations. (These orders of magnitude are substantially greater than those used in estimating an RfD or RfC in non cancer risk assessment). The NRC report explicitly considered variability only in the context of conversion at the dose level of observed range of animal risk data to human risk, without mention of the effect of extrapolation. These guidelines take the position that, given the sensitivity of the condition of test animals and the context of EPA practice, there is not a reason to add an uncertainty factor for human variation when a linear extrapolation procedure is used.

Margin of Exposure Analysis

There was a parallel call for more guidance and examples of margin of exposure analysis when mode of action indicates a nonlinear dose-response relationship.

More specific guidance for analysis of margin of exposure, including factors for inter-individual variability in toxicokinetic and toxicodynamic capacity (also see discussions of inter-individual variability above), will be developed, as recommended (SAB, 1999). These methodologies will be peer reviewed and published separately for application to health assessment generally.

Departing from Default Assumptions

A spectrum of views were expressed in the comments regarding the use of default assumptions (default inference options). They ranged from advocacy of invoking defaults only as a last resort to advocacy of never using any data to depart from a default unless the default can be disproved. The comments tended to focus primarily on the use of mode of action data. Little attention was given to assumptions such as the similarities of metabolic pathways among species that are more readily subject to definition through experimentation. Fears were expressed by different commenters that the Agency would be too lax and would be swayed by unfounded assertions from interested persons, or too strict and would ignore good science. No comments discussed the NRC, 1983 (Red Book) explanation of the use of inference options to support the continuation of the assessment process in the face of gaps in data or knowledge.

The Agency considers the use of defaults in the way that the Red Book outlined. They provide structure to the continuation of assessment in the face of gaps. All data sets for toxicity assessment are incomplete. Standard inferences must be in place as a consistent foundation for the everyday business of conducting a range of assessments for screening, for making priorities to gather exposure or toxicity data, and for in-depth examination to support risk management decisions. A completely unstructured system in which basic inferences are a last resort would be impracticable. A system that treats inferences as so rigid that they are virtually impossible to address with data would not be science. These final guidelines maintain the use of default assumptions for the purpose of providing structure and encouraging research as was proposed. To address the major concern that guidance was lacking on the issue of judging mode of action data, the guidelines contain a framework for assessing such data that has been subject to a special review by the Agency's Science Advisory Board (SAB, 1999). Moreover, the Agency will

continue to rely on scientific peer review to ensure that its scientific findings are sound and consistent with the current state of knowledge.

Peer Review

Comments on the proposal included several views about the intended use of peer review to support good science in Agency assessments and methods as recommended by the National Academy of Sciences (NRC, 1994). Most commenters were of the opinion that peer review is the right approach to make sure that risk assessments are sound and that controversial issues will be addressed with the help of the broader scientific community. The issue of composition of peer review panels was foremost in the comments of a minority of commenters. Opposing views were expressed in comments as to whether all Agency panels should exclude any scientist who works for industry. Comments were devoid of discussion whether anyone associated with any other groups should be excluded. Since proposing these guidelines, the Agency has adopted and published a Peer Review Handbook (U.S. EPA, 1998b). This Handbook contains Agency policy and practice on the technical products that should be peer reviewed, calls for diverse and balanced panels of experts from differing backgrounds and covers many other issues. The Handbook does not exclude otherwise qualified scientists from participating in peer review because of affiliation, but contains policies and practices to be followed to avoid compromises to the impartiality and independence of peer reviews. The reader is referred to the Handbook which contains extensive information and practices to be followed on the subject of affiliation and other issues regarding independent and impartial peer review.

[***Responses to other major issues (e.g. hazard descriptors, selection of the point of departure,) will be made available in the final guidelines.***]

_____ _____

Date

Carol M. Browner
Administrator

CONTENTS

LIST OF TABLES xxiv

LIST OF FIGURES xxv

1. INTRODUCTION 1-1
   1.1. PURPOSE AND SCOPE OF THE GUIDELINES 1-1
   1.2. ORGANIZATION AND APPLICATION OF THE GUIDELINES 1-2
      1.2.1. Organization 1-2
      1.2.2. Application 1-2
   1.3. USE OF DEFAULT ASSUMPTIONS 1-3
      1.3.1. Default Assumptions 1-4
      1.3.2. Major Defaults 1-10
   1.4. CHARACTERIZATIONS 1-19

2. HAZARD ASSESSMENT 2-1
   2.1. OVERVIEW OF HAZARD ASSESSMENT AND CHARACTERIZATION 2-1
      2.1.1. Analyses of Data 2-1
      2.1.2. Presentation of Results 2-1
   2.2. ANALYSIS OF TUMOR DATA 2-2
      2.2.1. Human Data 2-2
         2.2.1.1. Types of Studies 2-3
         2.2.1.2. Criteria for Assessing Adequacy of Epidemiologic Studies 2-4
         2.2.1.3. Criteria for Causality 2-7
         2.2.1.4. Assessment of Evidence of Carcinogenicity from Human Data . . . 2-9
      2.2.2. Animal Data 2-9
         2.2.2.1. Long-Term Carcinogenicity Studies 2-9
         2.2.2.2. Perinatal Carcinogenicity Studies 2-15
         2.2.2.3. Other Studies 2-16
      2.2.3. Structural Analogue Data 2-17
   2.3. ANALYSIS OF OTHER KEY DATA 2-17
      2.3.1. Physicochemical Properties 2-17
      2.3.2. Structure-Activity Relationships 2-18
      2.3.3. Comparative Metabolism and Toxicokinetics 2-19
      2.3.4. Toxicological and Clinical Findings 2-21
      2.3.5. Events Relevant to Mode of Carcinogenic Action 2-21
         2.3.5.1. Direct DNA Reactive Effects 2-21
         2.3.5.2. Indirect DNA Effects or Other Effects on Genes/Gene Expression 2-22
         2.3.5.3. Experimental Considerations in Evaluating Data on Precursor Events 2-24
         2.3.5.4. Judging Data 2-24

CONTENTS (continued)

2.4. BIOMARKER INFORMATION 2-25
2.5. MODE OF ACTION-GENERAL CONSIDERATIONS AND FRAMEWORK FOR ANALYSIS 2-26
2.5.1. General Considerations 2-26
2.5.2. Evaluating a Postulated Mode of Action 2-28
2.5.3. Framework for Evaluating a Postulated Carcinogenic Mode(s) of Action 2-29
2.5.3.1. Content of the Framework 2-30
2.6. WEIGHT-OF-EVIDENCE EVALUATION FOR POTENTIAL HUMAN CARCINOGENICITY 2-35
2.6.1. Weight-of-Evidence Analysis 2-35
2.6.2. Descriptors for Summarizing Weight of Evidence 2-44
2.7. TECHNICAL HAZARD CHARACTERIZATION 2-46
2.8. WEIGHT-OF-EVIDENCE NARRATIVE 2-48

3. DOSE-RESPONSE ASSESSMENT 3-1
3.1. HUMAN STUDIES 3-2
3.2. MODE OF ACTION AND DOSE-RESPONSE APPROACH 3-2
3.3. DOSE-RESPONSE ANALYSIS 3-4
3.3.1. Modeling the Overall Process--Biologically Based Models 3-4
3.3.2. Analysis in the Range of Observation 3-5
3.3.2.1. Applying Information About Key Events 3-5
3.3.2.2. Procedures for Analysis in the Range of Observation of Animal Studies 3-6
3.3.2.3. Point of Departure for Extrapolation from Observed Animal Data 3-6
3.3.3. Analysis in the Range of Extrapolation--Default Procedures 3-9
3.3.3.1. Linear Procedure 3-9
3.3.3.2. Nonlinear Extrapolation 3-10
3.3.3.3. Linear and Nonlinear Extrapolations 3-14
3.3.3.4. Use of Toxicity Equivalence Factors and Relative Potency Estimates 3-15
3.4. RESPONSE DATA 3-15
3.5. DOSE DATA 3-17
3.5.1. Interspecies Adjustment of Dose 3-18
3.5.2. Adjustment of Dose for Children 3-19
3.5.3. Toxicokinetic Analyses 3-20
3.5.4. Route-to-Route Extrapolation 3-21
3.5.5. Dose Averaging 3-22
3.6. DISCUSSION OF UNCERTAINTIES 3-23
3.7. TECHNICAL DOSE-RESPONSE CHARACTERIZATION 3-24

CONTENTS (continued)

4. TECHNICAL EXPOSURE CHARACTERIZATION 4-1

5. RISK CHARACTERIZATION 5-1
5.1. PURPOSE 5-1
5.2. APPLICATION 5-2
5.3. PRESENTATION OF RISK CHARACTERIZATION SUMMARY 5-3
5.4. CONTENT OF RISK CHARACTERIZATION SUMMARY 5-3
6. REFERENCES 6-1

APPENDIX A. WEIGHT-OF-EVIDENCE NARRATIVES A-1

APPENDIX B. RESPONSES TO THE NATIONAL ACADEMY OF SCIENCES NATIONAL RESEARCH COUNCIL REPORT *SCIENCE AND JUDGMENT IN RISK ASSESSMENT* (NRC, 1994) B-1

APPENDIX C. CASE STUDY EXAMPLES FOR HAZARD EVALUATION C-1

APPENDIX D. CASE STUDY EXAMPLES FOR MODE-OF-ACTION EVALUATION . D-1

APPENDIX E. NONLINEAR DOSE-RESPONSE: MARGIN-OF-EXPOSURE ANALYSIS (TO BE DEVELOPED) E-1

APPENDIX F. DOSE-RESPONSE ASSESSMENT FOR A CARCINOGEN POSING HIGHER RISKS AFTER CHILDHOOD EXPOSURE F-1

APPENDIX G. RESPONSE TO COMMENTS ON OTHER SCIENCE ISSUES (TO BE DEVELOPED) G-1

LIST OF TABLES

Table D-1. Thyroid follicular cell tumor incidence in male rats D-2
Table D-2. Incidence of transitional cell lesions and stones in males from a 2-year SD rat study D-12
Table D-3. Incidence of bladder hyperplasia and stones in male SD rats treated up to 13 weeks D-14
Table D-4. Reversal of incidence of bladder hyperplasia and stones following 8 weeks treatment and 16 weeks recovery D-16
Table D-5. Clinical chemistry values (urine) in male SD rats treated up to 13 weeks D-18
Table D-6. Summary results of chronic bioassays D-23
Table F-1. Comparison of tumor incidence in male and female Sprague-Dawley rats from 100 hour inhalation exposures to newborn and mature rats F-7
Table F-2 Comparison of tumor incidence in male and female Sprague-Dawley rats From 5-wk newborn exposure and 52-wk later life exposure.....F-8
Table F-3 Results of PBPK modeling.....F-9

LIST OF FIGURES

Risk characterization . . . . . .
Figure 2-1. Factors for weighing human evidence 2-36
Figure 2-2. 2-38
Factors for weighing other key evidence . . . . . .
Figure 2-4. Factors for weighing totality of evidence 2-42
Figure 3-1. 3-8